

IN THE SPECIFICATION:

Please amend the paragraph at page 1, line 23 through page 2, line 2 as follows:

--The present applicant has proposed a method of forming a hierarchical structure for each feature amount axis (also referred to as a dimension axis hereinafter) of a feature space and forming a classification tree on the basis of the hierarchical structure in Japanese Patent Application No. 9-187019 (U.S. Ser. Application No. 09/112,448, issued as U.S. Patent No. 6,560,359, issued May 6, 2003). --.

Please amend the paragraph at page 12, lines 16-25 as follows:

--For example, at the ~~route~~ root node, whether the learning patterns are classified at one of the two middle points ( $x = 0.5$  and  $y = 0.5$ ) (i.e., whether the X- or Y-axis hierarchical structure is used) must be determined. An index for this determination is, e.g., a "mutual information amount". This is the expected value of a decrease in entropy  $-\sum p \cdot \log(p)$  (see L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees, Chapman & Hall Inc., New York, N.Y., 1993, p. 33 for further detail). --.

Please amend the paragraph at page 13, lines 8-15 as follows:

--A practical process for selecting the dimension is shown with reference to FIG. 4. The category balance at the ~~route~~ root node is given as (A: 32, B: 32, entropy: 0.69). When the feature space is divided by the hyperplane of  $x = 0.5$ , the balance of the child node categories is given as (A: 5, B: 25, entropy: 0.45) in the range of  $x < 0.5$  and (A: 27, B: 7, entropy: 0.51) in the range of  $x \geq 0.5$ . --.

Please amend the paragraph starting at page 13, line 16 and ending at page 14, line 2 as follows:

--When the feature space is divided by the hyperplane of  $y = 0.5$ , the balance becomes (A: 20, B: 6, entropy: 0.57) in the range of  $y \geq 0.5$  and (A: 12, B: 26, entropy: 0.63) in the range of  $y < 0.5$ . These two choices are available at the ~~route~~ root node. In this case, the balance having a higher efficiency is selected using the entropy as the index for classification efficiency. In the former case, the expected value of the decrease in entropy is given as  $(30/64 \cdot 0.45 + 34/64 \cdot 0.51) - 0.69 = 0.21$ . In the latter case, the expected value is given as  $(26/64 \cdot 0.57 + 38/64 \cdot 0.63) - 0.69 = -0.08$ . Division of the learning patterns by the hyperplane of  $x = 0.5$  with a larger expected value of a decrease in entropy is selected. --.